

GRANT
IN-02-CR
164763
P.16

Final Technical Report
to the
Aircraft Guidance and Controls Branch/489
Guidance and Control Division
National Aeronautics and Space Administration
Langley Research Center
Hampton, VA 23665-5225

For the Grant NAG-1-1065

PARAMETER IDENTIFICATION FOR NONLINEAR AERODYNAMIC SYSTEMS

Period Covered: October 23, 1989 to February 28, 1993

from

Allan E. Pearson
Division of Engineering
Brown University
Providence, RI 02912

N93-27282

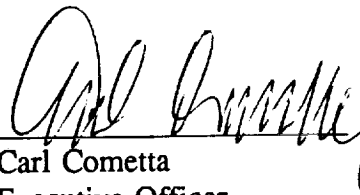
Unclass

G3/02 0164763

Report prepared by:



Allan E. Pearson
Professor of Engineering
Principal Investigator



Carl Cometta
Executive Officer
Division of Engineering

May 1993

(NASA-CR-193072) PARAMETER
IDENTIFICATION FOR NONLINEAR
AERODYNAMIC SYSTEMS Final Technical
Report, 23 Oct. 1989 - 28 Feb. 1993
(Brown Univ.) 16 p

Table of Contents

1.	Introduction	1
2.	List of Scientific Collaborators	1
3.	The System Identification Problem	1
4.	Research Results Under NAG-1-1065	2
4.1	The Explicitly Defined Fourier Modulating Functions	3
4.2	Least Squares Identification for L_g : SISO Case	5
4.3	Computational and Sampled-Data Considerations	6
4.4	The LS, WLS and AWLS Algorithms	9
4.5	Model Reduction	10
4.6	Extensions for L_g : MIMO Systems	10
4.7	Applications to an F-18 Aircraft	11
4.8	Extensions to LTV and Nonlinear System Models	11
4.9	Nonparametric Identification: Frequency Analysis	12
4.10	Miscellaneous Notes	12
5.	Publications and Reports	13

1. Introduction

This final technical report covers a three and one-half year period preceding February 28, 1993 during which support was provided under NASA Grant NAG-1-1065. Following a general description of the system identification problem and a brief survey of methods to attack it, the basic ideas behind the approach taken in this research effort are presented. The results obtained are described with reference to the published work (references [1-23] in Section 5), including the five semiannual progress reports previously submitted ([2,6,10,12,16]) and two interim technical reports ([7,17]).

2. List of Scientific Collaborators

Y. A. Fiagbedzi
A. V. Fullerton*
J. Q. Pan
A. A. Pandiscio
A. E. Pearson*
Y. Shen*

Assoc. Prof. of Math. Sci., King Fahd Univ.
Grad. Student Research Ass't. 1989-91
Grad. Student Research Ass't. 1989-92
Raytheon Grad. Student Fellow 1990-93
Professor and Principal Investigator
Grad. Student Research Ass't. 1991-93

* Received partial support under NAG-1-1065.

3. The System Identification Problem

The identification problem for a system with an input u and an output y is easy to state in generic terms: Given a class of systems with parametrized members S_θ such that an input/output pair (u, y) is modeled in the equation-error format $S_\theta(u, y) = 0$, find $\hat{\theta}$ such that a suitably chosen norm measure of distance, $\|S_{\hat{\theta}}(u, y)\|$, is "small" in some sense relative to all available i/o pairs (u, y) . The models considered in this research are those that can be described by input/output ordinary differential equations. In the general nonlinear system case, the parametrized members are represented implicitly by differential operator equations of the form:

$$S_\theta : \sum_{j=0}^{n_1} \sum_{k=1}^{n_2} g_j(\theta) F_{jk}(u, y) P_{jk}(p) E_k(u, y) = 0, \quad g_0 = 1 \quad (1)$$

where the $g_j(\theta)$ are specified functions of the parameters θ , the $P_{jk}(p)$ are specified polynomials of degree n in the differential operator $p = d/dt$, and the $(E_k(u, y), F_{jk}(u, y))$ are specified functions of the i/o pair (u, y) . Here the problem is to obtain an estimate of the parameters θ given sampled versions of the i/o data $[u(t), y(t)]$ over some time interval $0 \leq t \leq T$. Since parametrized state vector equations of the normal form:

$$\dot{x} = f(x, u, \theta), \quad y = h(x, u, \theta) \quad (2)$$

constitute the most common starting point for methods aimed at the parameter identification of nonlinear differential systems, the first step in our approach is to obtain (assuming this is possible) an equivalent i/o model like (1) when starting from a model of form (2). Methods for attacking the parameter identification of the model (2) include: quasilinearization, nonlinear filtering or smoothing, extended Kalman filtering, and state variable filters. The results of our efforts pertaining to the model (1) will be summarized in Section 4.8.

The more specialized but nevertheless omnipresent *linear* system counterparts to (1) and (2) are, respectively: the i/o differential operator models of the form ¹

$$L_{\theta} : \quad D(p)y(t) = N(p)u(t) \quad (3)$$

in which θ is comprised of the coefficients entering into the definitions of the polynomial matrix pair $(D(p), N(p))$, and the linear state vector equation model

$$\dot{x}(t) = Ax(t) + Bu(t), \quad y(t) = Cx(t). \quad (4)$$

In contrast to (1) and (2), the relations between (3) and (4) are well known, although the multi-input/multi-output case entails cumbersome details. Thus, if $G(s)$ denotes the transfer function matrix:

$$G(s) = D^{-1}(s)N(s) = C(sI - A)^{-1}B. \quad (5)$$

System identification for the class L_{θ} will also focus on determining the frequency transfer function $G(i\omega)$ given an i/o pair under transient conditions (Section 4.9).

Notwithstanding the problem of estimating the parameter vector θ , there is also the problem of *order* or *structure* determination for whatever class of models is under consideration. Effort in this regard has been confined to the linear class L_{θ} . In addition to the system identification approaches mentioned above for the general nonlinear model (2), the methods applicable to the models (3) and (4) include: approximation techniques like the δ operator, block pulse and orthogonal function expansion methods; the various moment functional techniques; and the conversion of an identified discrete-time model (obtained by a variety of methods) into a continuous-time model using, for example, the bilinear transformation.²

4. Research Results Under NAG-1-1065

The system identification research carried out under this grant has utilized the method of moment functionals which employs integration-by-parts to transfer derivatives on data variables to smooth user-chosen functions in order to obtain an algebraic equation useful for identification purposes. The particular method examined is originally due to Shinbrot ³ in which the user-chosen "modulating" functions possess the property that the resulting moment functionals are devoid of unknown initial/boundary conditions on the data, i.e., no state or initial condition estimation is involved for transient time-limited data. As introduced by Shinbrot, $\phi(t)$ is a *modulating function of order n* on a fixed time interval $0 \leq t \leq T$ if it is sufficiently smooth and satisfies the $2n$ end point conditions:

¹ Here we allow for multi-input/multi-output systems.

² Sources for discussion of these methods can be found in the recent books: *System Modeling and Identification*, by R. Johansson, Prentice Hall, 1993. *Identification of Continuous-Time Systems*, N. K. Sinha and G. P. Rao, eds., Kluwer Acad. Pub. 1991. *Identification of Continuous Systems*, by H. Unbehauen and G. P. Rao, North-Holland, 1987.

³ Shinbrot, M., "On the analysis of linear and nonlinear systems." *Trans. of the ASME*, 79, pp. 547-552, 1957. Also, NACA Rep. No. TN 3288, "On the analysis of linear and nonlinear systems from transient response data" by M. Shinbrot, 1954.

$$p^k \phi(t) = 0 \text{ at } t = 0 \text{ and } t = T, \quad k = 0, 1, \dots, n-1 \quad (6)$$

where p denotes the differential operator, i.e., $p = d/dt$, $p^2 = d^2/dt^2$, etc. Moreover, $\phi_m(t)$, $m=1, 2, \dots, M$, is a sequence of order n modulating functions if (6) is satisfied for each m .

Although Shinbrot's method and its unique property of avoiding state estimation has been known for a long time, it has remained relatively obscure due largely to a lack of clear guidance as to a good choice of modulating functions. In addition, the method may lack appeal on the basis of it being "off-line". As we reviewed in 1985,⁴ the 'cons' have apparently outweighed the 'pros' since very few papers have appeared on the method subsequent to Shinbrot's original work.⁵ What is needed is a sequence of modulating functions with highly composite members that ease the computational burden in calculating the underlying "moment functionals". The trigonometric Fourier modulating functions introduced in Pearson and Lee (Footnote 4) satisfied this need because the required functionals are Fourier series coefficients of the i/o data which can be calculated efficiently by DFT/FFT techniques. However, they were not presented in an explicit form, i.e., an *a priori* finite set of such modulating functions could be calculated by inverting a Vandermonde matrix depending on the order n and the number M of functions desired, but increasing n and/or M required a new inversion.⁶ This implicit stigma was removed by a discovery (see Eqn. (7) below) made during the early portion of the research carried out under NAG-1-1065. The explicit Fourier modulating functions facilitated several extensions of the theory and helped pave the way to more sophisticated methods for handling noise corrupted data.

4.1. The Explicitly Defined Fourier Modulating Functions

Consider the modulating function of order n defined by⁷

$$\phi_m(t) = \frac{1}{T} e^{-im\omega_0 t} (e^{-i\omega_0 T} - 1)^n, \quad 0 \leq t \leq T = 2\pi/\omega_0 \quad (7a)$$

$$= \frac{1}{T} \sum_{j=0}^n (-1)^{n-j} \binom{n}{j} e^{-i(m+j)\omega_0 t} \quad (7b)$$

where $i = \sqrt{-1}$, $\omega_0 = 2\pi/T$ plays the role of a 'resolving' frequency, and m is any integer which shall be referred to as the 'modulating frequency index'. The fact that $\phi_m(t)$ is a modulating function of order n for each m is evident from the basic definition (6) applied to (7a). Thus,

⁴ A. E. Pearson and F. C. Lee, "Parameter identification of linear differential systems via Fourier based modulating functions," *Control-Theory and Adv. Tech.*, **1**, pp. 239-266, 1985.

⁵ This remains true today. In fact, the recent book edited by Sinha and Rao (see Footnote 2), although a good source for a variety of topics dealing with differential system identification, contains no reference to Shinbrot's method.

⁶ In the paper: "System identification using modulating functions and fast Fourier transforms," by T. B. Co and B. E. Ydstie, *Computer Chem. Engr.*, **14**, pp. 1051-1056, 1990, independent verification has been given in the use of the real Fourier modulating functions by application to a variety of ODE models, as well as showing how the method can be used in model reduction.

⁷ Equation (7b) follows from (7a) upon utilizing the binomial expansion.

the ϕ_m in (7), $m=0,1,\dots,M$, represent an explicitly defined linearly independent set of order n modulating functions with the important property that moment functionals involve Fourier series coefficients on $[0,T]$.

Two modulation properties can be developed that ease the various least squares formulations. The first concerns a linear differential operator $P(p)$, i.e., a polynomial like $\sum_{k=0}^n a_k p^{n-k}$, and a function $v(t)$ on $[0,T]$ with Fourier series coefficients $V(m)$ defined by

$$V(m) = \frac{1}{T} \int_0^T v(t) e^{-im\omega_0 t} dt. \quad (8)$$

Property 1. For a sufficiently smooth function $v(t)$ defined on $[0,T]$ and a differential operator $P(p)$ of order n , the integration of $\phi_m(t)P(p)v(t)$ over $[0,T]$ satisfies

$$\int_0^T \phi_m(t) P(p) v(t) dt = \Delta^n P(im\omega_0) V(m) \quad (9)$$

where Δ^n is the n^{th} order finite difference operator, i.e., with $Q(m) = P(im\omega_0)V(m)$,

$$\Delta Q(m) = Q(m+1) - Q(m)$$

$$\Delta^2 Q(m) = Q(m+2) - 2Q(m+1) + Q(m)$$

$$\Delta^n Q(m) = \sum_{k=0}^n (-1)^k \binom{n}{k} Q(n+m-k). \quad (10)$$

Proof. Taking into account (6) and using integration-by-parts on the left side of (9), all boundary point evaluations are zeroed while transferring the derivatives on $v(t)$ to derivatives on the modulating function $\phi_m(t)$ such that:

$$\int_0^T \phi_m(t) P(p) v(t) dt = \int_0^T v(t) P(-p) \phi_m(t) dt.$$

In view of the representation (7b), coupled with the fact that $P(-p)e^{-\lambda t} = P(\lambda)e^{-\lambda t}$, the right sides of (9) and (10) are obtained which proves the property.

The implication of (9) is that derivatives of signals in the continuous time domain can be traded exactly for finite differences of modulated Fourier series coefficients in the discrete frequency domain by forming moment functionals using the Fourier modulating functions. The presence of such Fourier series coefficients should not be interpreted as implying a kind of Fourier series approximation in the identification problem. Rather, Property 1 facilitates an equivalent statement of the system identification problem entirely in the frequency domain while utilizing transient i/o data. In applications, the range of values to be chosen for m and the choice of the time interval $[0,T]$, equivalently the resolving frequency ω_0 , will depend on the bandwidth of the signals involved, as well as the complexity of the model relating the signals. It is noted that $\Delta^n X$ can be calculated by the M-file 'diff(X,n)' in MATLAB, which facilitates ease of coding in that environment, and the Fourier series coefficients of the signals required in the identification problem can be calculated efficiently and accurately via

DFT/FFT techniques. Further elaboration on this point follows below in Section 4.3.

4.2. Least Squares Identification for L_θ : SISO Case

Let an i/o data pair $[u(t), y(t)]$ for a single-input/single-output linear system of order n be modeled on $[0, T]$ by

$$p^n y(t) + \sum_{j=1}^n a_j p^{n-j} y(t) = \sum_{j=1}^n b_j p^{n-j} u(t) + e(t) \quad (11)$$

where the (a_j, b_j) , $j=1, 2, \dots, n$, are parameters and $e(t)$ represents the effect of modeling errors. Modulating (11) with $\phi_m(t)$, integrating over $[0, T]$ and applying Property 1:

$$\Delta^n(im\omega_0)^n Y(m) + \sum_{j=1}^n a_j \Delta^n(im\omega_0)^{n-j} Y(m) = \sum_{j=1}^n b_j \Delta^n(im\omega_0)^{n-j} U(m) + \varepsilon(m) \quad (12)$$

where the residuals in (12) are defined by

$$\varepsilon(m) = \int_0^T e(t) \phi_m(t) dt. \quad (13)$$

Defining a column vector $\theta = \text{col}[-a_1, \dots, -a_n, b_1, \dots, b_n]$, a rearrangement of terms in (12) leads to the standard (albeit complex-valued) linear regression:

$$\gamma_0^y(m) = \gamma(m) \theta + \varepsilon(m), \quad m = 0, 1, \dots, M \quad (14)$$

where the row vector of regressors $\gamma(m)$ is defined by

$$\gamma(m) = \text{row}(\gamma_1^y(m), \dots, \gamma_n^y(m), \gamma_1^u(m), \dots, \gamma_n^u(m)) \quad (15)$$

and the pairs $(\gamma_j^u(m), \gamma_j^y(m))$, $j=0, 1, \dots, n$, are defined by

$$\begin{bmatrix} \gamma_j^u(m) \\ \gamma_j^y(m) \end{bmatrix} = \Delta^n(im\omega_0)^{n-j} \begin{bmatrix} U(m) \\ Y(m) \end{bmatrix}. \quad (16)$$

Taking into account the fact that (14) is complex-valued, a cost function for least squares minimization can be defined by

$$J(\theta) = (Y_c - \Gamma_c \theta)' (Y_c - \Gamma_c \theta) \quad (17)$$

where prime denotes transpose and the following notation applies:

$$Y_c = \begin{bmatrix} \text{Re } \gamma_0^y(0) \\ \vdots \\ \text{Re } \gamma_0^y(M) \\ \text{Im } \gamma_0^y(0) \\ \vdots \\ \text{Im } \gamma_0^y(M) \end{bmatrix}, \quad \Gamma_c = \begin{bmatrix} \text{Re } \gamma(0) \\ \vdots \\ \text{Re } \gamma(M) \\ \text{Im } \gamma(0) \\ \vdots \\ \text{Im } \gamma(M) \end{bmatrix} \quad (18)$$

with the complex $\gamma(m)$ and $\gamma_0^y(m)$ defined in (15) and (16).

Assuming linearly independent regressors, the one-shot estimate which minimizes (17) is given by

$$\hat{\theta} = (\Gamma_c' \Gamma_c)^{-1} \Gamma_c' Y_c. \quad (19)$$

Noting from (10) that M in (18) actually corresponds to the frequency $(M+n)\omega_0$ suggests that in choosing the pair (M, ω_0) , their product should correspond roughly to the system bandwidth. Also, M has to be sufficiently large to provide enough algebraic equations upon which to base a one-shot estimate. Hence, if $\#$ denotes the number of unknowns ($\# = 2n$ for the model (11)), then the pair (M, ω_0) can be uniquely determined by the guidelines:

$$M \approx 2\# - 4\#$$

$$\text{and} \quad (M+n)\omega_0 \approx \omega_B \quad (20)$$

where ω_B is the system bandwidth, i.e., the highest useful frequency in the i/o data.

4.3. Computational and Sampled-Data Considerations

Assuming uniform sampling of the i/o data on $[0, T]$, the integral (8) for all required Fourier series coefficients will be approximated by standard quadrature formulae which, in turn, are calculated by a DFT/FFT algorithm. Thus, if the sampling interval δt and the total number of samples N satisfy $\delta t = T/N$ and $N > M+n$, the staircase, trapezoidal and parabolic rules applied to (8) lead respectively to

$$\frac{1}{T} \int_0^T v(t) e^{-im\omega_0 t} dt = \frac{1}{N} \left[\sum_{k=0}^{N-1} v_k W^{mk} \right] + O\left(\frac{1}{N}\right) \quad (21a)$$

$$\frac{1}{T} \int_0^T v(t) e^{-im\omega_0 t} dt = \frac{1}{N} \left[\frac{v_0 + v_N}{2} + \sum_{k=1}^{N-1} v_k W^{mk} \right] + O\left(\frac{1}{N^2}\right) \quad (21b)$$

$$\frac{1}{T} \int_0^T v(t) e^{-im\omega_0 t} dt = \frac{2}{3N} \left[\frac{v_0 + v_N}{2} + \sum_{k=1,3,\dots}^{N-1} 2v_k W^{mk} + \sum_{k=2,4,\dots}^{N-1} v_k W^{mk} \right] + O\left(\frac{1}{N^4}\right) \quad (21c)$$

where

$$m = 0, 1, \dots, M+n, \quad v_k = v(kT/N) \quad \text{and} \quad W = e^{-i2\pi/N}.$$

The quantities within brackets on the right hand sides of the above can be evaluated by a standard DFT/FFT algorithm provided the range of values for the modulating frequency index m is artificially extended to $[0, N-1]$. A fundamental property of the DFT is that the first and last halves of the transformed sequence are complex conjugates of one another. Hence, in order to preserve uniqueness for the first $M+n$ harmonics, it follows that N has to satisfy

$$N/2 \geq M + n. \quad (22)$$

In practical experience using either simulated or physical data, very little difference is noticed between the three approximations in (21) provided: (i) the data is oversampled in terms of the Nyquist sampling theorem, and (ii) the guideline (20) is upheld.⁸ These provisos

⁸ The first proviso is more essential than the second if the "adaptive" weighted least squares (AWLS) algorithm is utilized (Section 4.4), i.e., the AWLS algorithm tends to make the parameter estimates insensitive to a mismatch in bandwidth.

also bear a relation to the DFT requirement (22) as can be shown by the following. Let F_B and F_S denote the bandwidth and sampling frequency respectively (in Hertz), then 'oversampling' implies the inequality:

$$2F_B \ll F_S \Rightarrow \frac{\omega_B}{\pi} \ll \frac{1}{\delta t} = \frac{N}{T} = \frac{N\omega_0}{2\pi}.$$

Assuming adherence of the bandwidth guideline (20), the above implies

$$\frac{(M+n)\omega_0}{\pi} \ll \frac{N\omega_0}{2\pi} \Rightarrow \frac{M+n}{N/2} \ll 1. \quad (23)$$

Hence, the DFT requirement (22) is amply upheld by oversampling and adherence to the bandwidth guideline in (20); moreover, (23) implies that only the lowest few percent of the available DFT harmonics are utilized in the least squares identification.

The following example from [22] illustrates some aspects of the sampling rate consideration and, at the same time, provides a comparison with a commercially available algorithm for parameter identification. Consider the system

$$\ddot{y}(t) + 3\dot{y}(t) + 8y(t) = 5u(t), \quad 0 \leq t \leq T = 10s \quad (24)$$

which possesses the transfer function: $H(s) = 5/(s^2 + 3s + 8)$. Shown in Fig. 1 for point of reference is: (a) the time domain step response of the system, (b) the frequency domain magnitude plot, and (c) the effect of sampling rate on the parameter estimation errors under ideal zero-noise conditions. The latter is discussed below. Since there are only 3 unknown parameters and $n=2$, M was chosen as $M=6$. This means that $(M+n)\omega_0=0.8$ Hz is the highest frequency that will be extracted from the i/o data, which is roughly the bandwidth of the system as seen in Fig. 1(b).

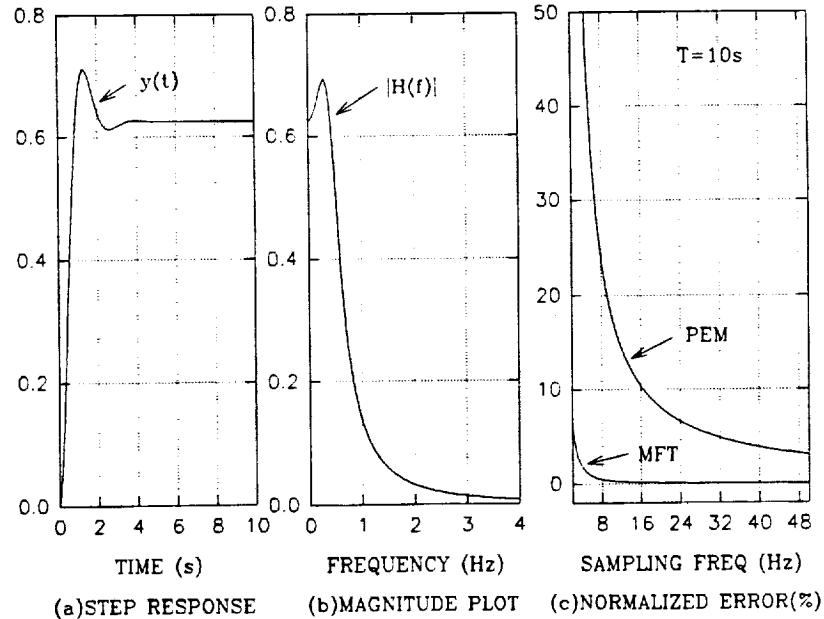


Fig. 1. Aspects of the simulated system

The output $y(t)$ was corrupted by additive white Gaussian measurement noise. Two hundred Monte Carlo runs were made for each of several noise-to-signal ratios (NSR) under two separate conditions: (i) the initial conditions fixed at $(0,0)$, and (ii) the initial conditions randomized for each run, i.e., a total of 400 Monte Carlo runs at each NSR. The input was $u(t)=\sin^2 t/5$ over the $T=10s$ time interval for each run. All calculations were carried out in MATLAB. The sampling rate was fixed at 25.6 Hz, thus facilitating a standard 256 point DFT for the 10s time interval, and the staircase rule (21a) was used in approximating (8) to obtain the i/o Fourier series coefficient pairs $[U(m), Y(m)]$, $m=0,1 \cdots M+n$. Since $N=256$ is much larger than $M+n=8$, it is noted from (23) that only $8/128 \approx 6\%$ of the total harmonics available from the i/o DFT's is actually retained in implementing the algorithm. This accounts for the high accuracy achieved in implementing the method with modest sampling rates inasmuch as calculating the higher frequency coefficients would be expected to entail greater numerical error in the presence of high frequency noise. This is further discussed below.

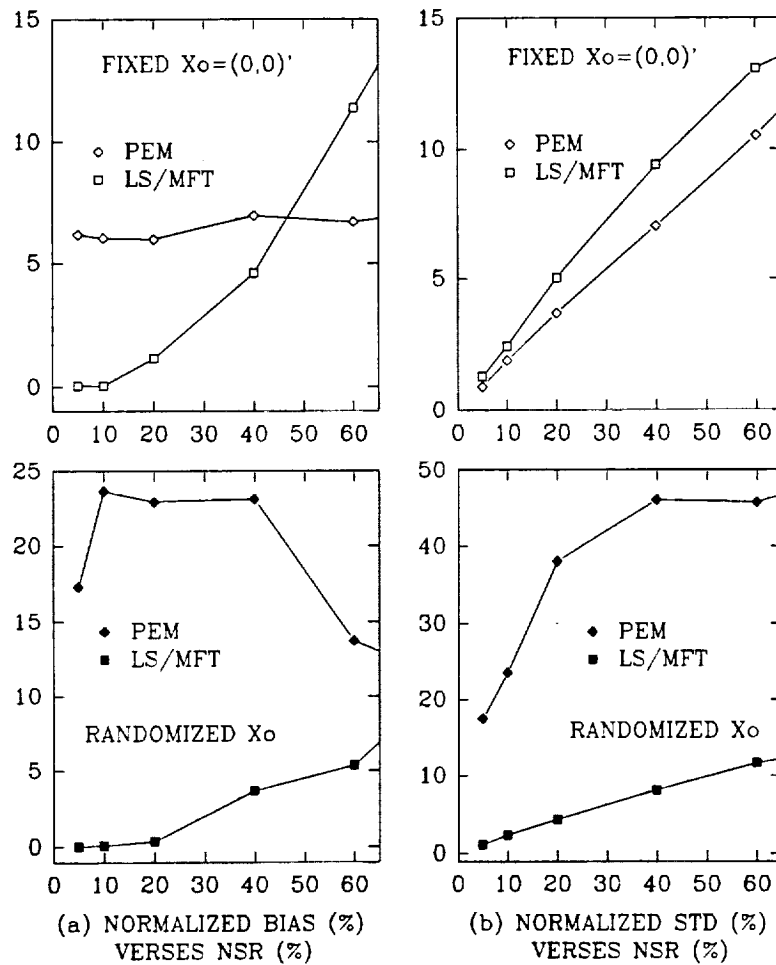


Fig. 2. Measurement noise effects

The results are shown in Fig. 2 where "MFT" means the estimate is based on the modulating function technique. Also shown is the estimate based on the prediction error method (PEM) from the Identification Toolbox in MATLAB.⁹ In the case of nonrandom (fixed) initial conditions shown in the top halves of Fig. 2, the PEM shows a fairly constant 6~7% normalized bias error¹⁰ as the NSR increases from 0% to 60%, whereas the bias for the MFT gradually increases from 0% to ~12% over the same range of NSR values. Meanwhile, the normalized standard deviations are nearly the same for this case. However, in the case of randomized initial conditions, see the lower halves of Fig. 2, both the bias and the standard deviation for the PEM show a several-fold increase and, moreover, each exhibits rather unpredictable behavior that is difficult to reproduce even with additional Monte Carlo runs. By contrast, the bias and standard deviation have both decreased at each NSR for the MFT, indicating that nonzero initial conditions favor the MFT.

The above comparative results have been verified on other examples as well (see Fullerton [11]). One reason for this is that the MFT does not have to estimate unknown initial conditions; another is related to the fact that the MFT is a direct identification technique for continuous-time models, while the PEM first estimates the parameters of a discrete-time model then converts this to a continuous-time model. Also, the Fourier series coefficients needed by the MFT can be computed accurately with modest sampling rates using the DFT. Additional insight in this regard is gained by referring to Fig. 1(c) which shows the influence of sampling rate on the percent normalized error under zero measurement noise conditions. Apparently the PEM needs a much higher sampling rate for the given 10s of data, or a much longer T time interval for the given 25.6 Hz sampling rate, in order to achieve a small estimation error. The influence of sampling rate on the estimation error for the MFT is much more consistent with the Nyquist sampling theorem in view of the frequency magnitude plot in Fig. 1(b). Finally, it is interesting to note that the MFT requires substantially less computer time for each set of Monte Carlo runs (about a ten to five fold decrease depending on whether or not the PEM has to estimate the initial conditions).

4.4. The LS, WLS and AWLS Algorithms

If the equation error function $e(t)$ in the linear system model (11) is modeled as stationary Gaussian "white" noise with zero mean and covariance $Ee(t)e(t+\tau)=\sigma\delta(\tau)$, then the modulated residual process $\epsilon(m)$ defined in (13) is also Gaussian with a covariance matrix given by Equation (8) in [16], or (2.25) in [20]. Even though the residuals in the identification problem will not be white, this fact facilitates a reasonable weighting matrix in the discrete frequency domain for a weighted least squares estimate which attempts to ameliorate the effects of noise when the measurement noises on the i/o data are white. Moreover, it points the way to formulating an "adaptive" weighted least squares algorithm in the

⁹ Ljung, L. (1987). *System Identification Toolbox*, Version 2.11. The MathWorks, Inc.

¹⁰ Defined as $\left[\frac{1}{\#} \sum_{j=1}^{\#} \left[\frac{\hat{\theta}_j - \theta_j^*}{\theta_j^*} \right]^2 \right]^{\frac{1}{2}} \cdot 100\%$ where $\hat{\theta}_j$ is the ensemble average of the MFT estimates for the true θ_j^* .

discrete frequency domain which emulates an approximate maximum likelihood estimate. These are the basic ideas developed in Chapter 2 of Shen's Ph.D. thesis [20]. The advantage of these developments is illustrated by a simulation study carried out for the system (24) and summarized in Section 2.4 of [20]. In addition to the sampled i/o data on $[0, T]$, implementation of the AWLS algorithm requires an assumption of the system order n and an initial guess for the weighting matrix W which is normally taken as the covariance matrix under the ideal white noise residual case. Convergence is fast and has never failed in numerous trials. The software for these algorithms will be supplied under separate cover in a toolbox to be prepared by Y. Shen and the PI. An earlier toolbox based on the real Fourier modulating functions was prepared by Fullerton [8].

4.5. Model Reduction

Motivated in part by the requirements attending one of the two methods developed in this research to handle MIMO models, the AWLS algorithm was applied to the model reduction problem, thus giving a frequency domain approach to this problem that has several advantages: (i) the high order model to be simplified can be supplied in either the time domain via a differential equation, or in the frequency domain via Bode type plots, (ii) weighting the cost function in the frequency domain is a straightforward and direct option for the AWLS formulation, and (iii) the resulting lower order model obtained via the AWLS formulation is at least as good and usually better than the same order model obtained by all methods with which the algorithm has been compared. Most notably, this is true for the well-known Balanced Reduction method, which is available in MATLAB. The detailed discussion and verification of this is given in Chapter 3 of Shen's thesis [20].

4.6. Extensions for L_θ : MIMO Systems

Two viewpoints in modeling multi-input/multi-output systems led to two different formulations to extend the LS, WLS and AWLS algorithms for the class L_θ in (3). If the underlying model is selected to be a state space model as in (4), then an appropriate scalar-valued i/o differential operator model, which is required in this approach, is the following:

$$P(p)y_j(t) = \sum_{k=1}^{m_u} Q_{jk}(p)u_k(t), \quad 0 \leq t \leq T \quad (25)$$

$$j = 1, 2, \dots, m_y$$

where the scalar-valued polynomial $P(s)$ is defined by $P(s) = \det(sI - A)$ in relation to the system matrix A in (4). A joint cost function that reflects the equation errors for each equation in (25) can be specified and minimized using the AWLS algorithm. A weighting between each such equation error has to be supplied, or determined by testing different weightings, which essentially accounts for differences in the output measurement noise variances for the m_y outputs. This is the viewpoint taken in Chapter 5 of Shen's thesis [20] (also summarized in [16]) which was used to model the physical data supplied under NAG-1-1065 for an F-18 aircraft.

Another viewpoint in modeling MIMO systems of the class L_θ is to find the best linear transfer function model that relates the i/o data without constraining the polynomials on the left hand side of (25) to be the same for each j . That is, each output y_j could be modeled

by:

$$P_j(p)y_j(t) = \sum_{k=1}^{m_u} Q_{jk}(p)u_k(t), \quad 0 \leq t \leq T \quad (26)$$

which differs from (25) in that the poles of the estimated MISO models are not constrained to be the same. This leads to a simpler extension of the AWLS algorithm developed for the SISO case, but necessitates a model reduction for each resulting MISO model in (26) in order to ferret out the common poles and zeros. This viewpoint is developed in Chapter 4 of Shen's thesis [20] and motivated the model reduction results mentioned in Section 4.2 above.

4.7. Applications to an F-18 Aircraft

Using physical data for an F-18 aircraft supplied under NAG-1-1065, the AWLS algorithm was applied to obtain i/o models for both the lateral and longitudinal dynamics [16,17,20]. The following output signal-to-error ratio (in decibels) was used as a measure of goodness-of-fit in each channel:

$$S/E = 20 \log_{10} \frac{\text{RMS}(y)}{\text{RMS}(e)}, \quad e(t) = y(t) - \hat{y}(t), \quad 0 \leq t \leq T$$

where $\text{RMS}(y)$ means the root-mean-square value of the output signal $y(t)$ over the $[0, T]$ time interval, and $\hat{y}(t)$ is the estimated output using the model. Since our algorithm does not estimate initial conditions, we used a standard Luenberger observer running backwards in time (see details in Appendix C of Shen [20]) in order to estimate an initial condition and, subsequently, simulate the model with the given input data to obtain $e(t)$ for the above S/E . Typical S/E ratios were on the order of 8 ~ 12 db for the longitudinal dynamics, depending on whether a second or third order model is used, and 9 ~ 12 db for the lateral dynamics. (See Figs. 5.5 and 5.12 in Shen [20] for the time domain performance of the resulting models.) In order to compare with a simulated model under specified measurement noise conditions, the same algorithm was applied to a theoretical linear model supplied to us under NAG-1-1065 using the given forcing function data for the physical modeling. These models gave S/E ratios on the order of 34 ~ 41 db for the longitudinal dynamics and 16 ~ 30 db for the lateral dynamic channels (see Figs. 5.8 and 5.14 in [20] for the time domain performance of these models). Even at 16 db the fit is very good, and at S/E ratios above 30 db there is hardly any discernible error in the model responses compared with the theoretical outputs.

4.8. Extensions to LTV and Nonlinear System Models

If the time variations in the coefficients of a linear time-varying differential equation model can be parametrized as linear combinations of known modal (smooth) functions on $[0, T]$, then the least squares formulation developed in Section 4.2 can be extended to yield modulated equation errors in the discrete frequency domain with Fourier series coefficients computed for time-modulated i/o data, i.e., the i/o data modulated by the modal functions on $[0, T]$. This is carried out in [22] and is illustrated by simulations for a second order system with a parabolically varying damping coefficient. The good noise rejection properties experienced with the algorithm for time-invariant systems appear to carry over to the time-varying case as well.

If the form of the nonlinearities are known up to parametric representation and if the i/o model can be manipulated into the generic form (1), then it is possible to extend the least squares parameter identification in the discrete frequency domain. This is carried out in a series of publications in [1,18,19]. Although mainly of theoretical interest thus far, the potential exists to apply these algorithms to unsteady aerodynamic models akin to those studied in Chin and Lan.¹¹

4.9. Nonparametric Identification: Frequency Analysis

Methods for determining the frequency transfer function of a stable linear system from i/o data include correlation and spectral analyses, as well as the direct ratio of Fourier transforms and steady state sinusoidal measurements. Each of these "nonparametric" identification techniques require either a statistical stationarity assumption on the data, or a periodic steady state condition to be established, before initiating calculations of the transfer function at pertinent frequencies. Notwithstanding noise considerations, long data lengths may be required in order to achieve good accuracy due to the stationarity or steady state assumption, or to eliminate end point effects in case a direct ratio of Fourier transforms is used on time-limited data. By contrast, a method was proposed in [13,14] that utilizes the modulating functions (7) to extract the frequency content in short data lengths in order to set up a least squares estimation of the transfer function at selected frequencies. Since short data lengths are used there is no assumption of steady state operation or stationarity of the data, though there must be present sufficient energy content in the data at the specified frequencies in order to avoid degeneracy in the least squares estimate. The method was also extended to multivariable systems, and was illustrated by way of simulation for a two-input/two-output example in [22].

4.10. Miscellaneous Notes

In addition to the results summarized above, research under NAG-1-1065 has included the following topics.

Order Determination for L_θ The problem of determining the orders of the various polynomials that enter into specifying a member of the linear class L_θ has been studied in [9,14,20]. Basically, the approach taken is to examine the residuals $\epsilon(m)$ in the discrete frequency domain for various presumed orders starting from a lowest order. Since there are no initial conditions to be estimated in the modulating function approach, the norm of the residuals in the frequency domain is unbiased by unknown initial/boundary conditions and provides a useful measure for order determination. However, more could probably be done on this important problem and, hence, this remains a topic for future investigation.

Collinearity The problem of collinearity in the regressors brought about by linear feedback was studied in [11]. The result is that predicting the deleterious effects of noise and collinearity remains a difficult problem in system identification, but the degree of collinearity that can

¹¹ Chin, S. and C. E. Lan (1991). *Fourier Functional Analysis for Unsteady Aerodynamic Modeling*. AIAA-91-2867-CP.

be tolerated in the modulating function approach is quite high, e.g., about 95% degree collinearity (as defined in [11]) while maintaining about 5% accuracy in the estimated parameters and in the presence of about 5% noise on the data.

I/O Modeling to Alleviate Noisy Measurements A study was undertaken in [7] to derive an i/o model not involving a noisy velocity signal with the expectation that the accuracy in the parameter estimates might be enhanced if the signal in question was particularly noisy. The state equation model with appropriate noise levels was communicated to us from E. A. Morelli.¹² The expectation was realized for those parameters that could be estimated from the i/o model, i.e., not all the state equation parameters could be estimated uniquely from the i/o model due to an inherent lack of identifiability for certain state equation parameters in the input/output model.

Frequency Estimation in a Signal Processing Problem The classic problem of estimating the frequencies in a sum of sinusoids buried in noise can be formulated as a system identification problem of the class L_0 using the fact that the sinusoids satisfy a homogeneous linear differential equation. The modulating function approach was used to study this problem with the expectation that the avoidance of initial condition estimation could enhance the accuracy in the estimates. The results of the study are contained in [14,15] and basically confirm this expectation in relation to the high order Yule-Walker approach.

Time Lag Systems Although not funded by NAG-1-1065, research results have been obtained on an approach to the feedback control of time lag systems that employs a "reducing transformation" [3-5,21]. The main advantage of this approach is that it facilitates the design of feedback controllers for systems described by differential-delay equations based on well-known finite dimensional control techniques. Our recent studies in [21] extend robust control techniques, specifically the structured singular value, to this class of control systems within the context of the reducing transformation method.

5. Publications and Reports

- [1] Pearson, A. E., "Identifiability and Well-Posedness in Nonlinear System I/O Modeling," *Proc. of 28th IEEE CDC*, pp. 624-625, Tampa, FL, Dec. 1989.
- [2] Semiannual Progress Report No. 1 for NAG-1-1065, A. E. Pearson-PI, April 26, 1990.
- [3] Fiagbedzi, Y. A. and A. E. Pearson, "Exponential State Observer for Time Lag Systems," *Int. J. Control*, **51**, no. 1, pp. 189-204, 1990.
- [4] Fiagbedzi, Y. A. and A. E. Pearson, "Output Feedback Stabilization of Delay Systems Via Generalization of the Transformation Method," *Int. J. Control*, vol. 51, no. 4, pp. 801-822, 1990.
- [5] Fiagbedzi, Y. A. and A. E. Pearson, "A State Observer for Systems Described by Functional Differential Equations," *Automatica*, **26**, no. 2, pp. 321-331, March 1990.

¹² Morelli, E. A., "Practical Input Optimization for Aircraft Parameter Estimation Experiments," Sc.D. Dissertation, JIAFS, George Washington Univ., Washington, DC, 1990.

- [6] Semiannual Progress Report No. 2 for NAG-1-1065, A. E. Pearson - PI, Nov. 19, 1990.
- [7] On Parameter Identification for an F-18 Aircraft, Interim Technical Report for NAG-1-1065, A. E. Pearson - PI, January 29, 1991.
- [8] Fullerton, A., "PLF Toolbox User's Guide: Modulating Function Method for use with MATLAB," Version 2.1, Div. of Engr., Brown Univ., March 1991.
- [9] Pearson, A. E. and J. Q. Pan, "On Order Determination for Linear Differential Scalar Systems," *Proc. of 25th CISS*, pg. 161, Johns Hopkins U., Baltimore, MD, March 1991.
- [10] Semiannual Progress Report No. 3 for NAG-1-1065, A. E. Pearson - PI, May 20, 1991.
- [11] Fullerton, A., "Noise and Collinearity in Differential Systems Identification," Sc.M. Thesis, Div. of Engr., Brown Univ., Providence, RI, May 1991.
- [12] Semiannual Progress Report No. 4 for NAG-1-1065, A. E. Pearson - PI, October 30, 1991.
- [13] Pearson, A. E. and J. Q. Pan, "Frequency Analysis Via The Method of Moment Functionals," *Proc. of 30th IEEE CDC*, pp. 2024-2025, Brighton, UK, Dec. 1991.
- [14] Pan, J. Q., "System Identification, Model Reduction and Deconvolution Filtering Using Fourier Based Modulating Signals and High Order Statistics," Ph.D. Thesis, Brown Univ., Providence, RI, Feb. 1992. Also, Technical Report LEMS-103, Division of Engineering, Brown Univ., March 1992.
- [15] Pan, J. Q. and A. E. Pearson, "High Resolution Frequency Estimation in the Presence of Noise Using Complex Sinusoidal Modulating Signals," *Proc. of 1992 CISS*, pg. 432, Princeton Univ., Princeton, NJ, March 1992.
- [16] Semiannual Progress Report No. 5 for NAG-1-1065, A. E. Pearson - PI, May 7, 1992.
- [17] *Parameter Identification and Model Comparisons for the F-18 Trial d Data*, Interim Technical Report for NAG-1-1065, A. E. Pearson - PI, August 14, 1992.
- [18] Pearson, A. E., "Explicit Parameter Identification for a Class of Nonlinear Input/Output Differential Operator Models," *Proc. of 31st IEEE CDC*, pp. 3656-3660, Tucson, AZ, Dec. 1992.
- [19] Pearson, A. E., "Explicit Least Squares System Parameter Identification for Exact Differential Input/Output Models," *Proc. of the Eighth ICMCM*, X. J. R. Avula, ed., in *Math. Modelling and Sci. Computing*, 2, pp. 101-107, 1993.
- [20] Shen, Y., "System Identification and Model Reduction Using Modulating Function Techniques," Ph.D. Thesis, Brown Univ., Providence, RI, Feb. 1993. Also, Technical Report LEMS-117, Division of Engineering, Brown Univ., March 1993.
- [21] Pandiscio, A. A. and A. E. Pearson, "The Reducing Transformation and Smith Predictor Methods for Stabilizing Plants with Transport Lag." Invited paper to appear in the Proceedings of the 1993 ACC.
- [22] Pearson, A. E., Y. Shen and J. Q. Pan, "Discrete Frequency Formats for Linear Differential System Identification." To appear in the Proceedings of the 1993 IFAC Congress.
- [23] Pearson, A. E. and Y. Shen, "Weighted Least Squares/MFT Algorithms for Linear Differential System Identification." Submitted to the 1993 IEEE CDC.